QoLAS: A Reddit Corpus of Health-Related Quality of Life Aspects of Mental Disorders

Lynn Greschner[1], **Amelie Wührl**[3,1,2], and **Roman Klinger**[1]
[1]Fundamentals of Natural Language Processing, University of Bamberg
[2]Institute for Natural Language Processing, University of Stuttgart
[3]IT University of Copenhagen
{lynn.greschner,roman.klinger}@uni-bamberg.de, amwy@itu.dk

Abstract

'Quality of Life' (QoL) refers to a person's subjective perception of various aspects of their life. For medical practitioners, it is one of the most important concepts for treatment decisions. Therefore, it is essential to understand in which aspects a medical condition affects a patient's subjective perception of their life. With this paper, we focus on the under-resourced domain of mental health-related QoL, and contribute the first corpus to study and model this concept: We (1) annotate 240 Reddit posts with a set of 11 QoL aspects (such as 'independence', 'mood', or 'relationships') and their sentiment polarity. Based on this novel corpus, we (2) evaluate a pipeline to detect QoL mentions and classify them into aspects using open-domain aspect-based sentiment analysis. We find that users frequently discuss health-related QoL in their posts, focusing primarily on the aspects 'relationships' and 'selfimage'. Our method reliably predicts such mentions and their sentiment, however, detecting fine-grained individual aspects remains challenging. An analysis of a large corpus of automatically labeled data reveals that social media content contains novel aspects pertinent to patients that are not covered by existing QoL taxonomies.

1 Introduction

'Quality of Life' (QoL) refers to a person's subjective perception considering various aspects of their life (World Health Organization, 2012). In the medical domain, understanding individual QoL aspects is crucial as they determine appropriate treatments for patients. Traditionally, QoL is assessed by medical experts, for instance, with the help of questionnaires in a personal interaction with a patient. While this approach benefits individuals with access to healthcare, these small-scale assessments are expensive, limited regarding individual repercussions of a medical condition, and are potentially subject to reporting biases. Therefore, this approach is not ideal for large-scale analysis of QoL aspects.

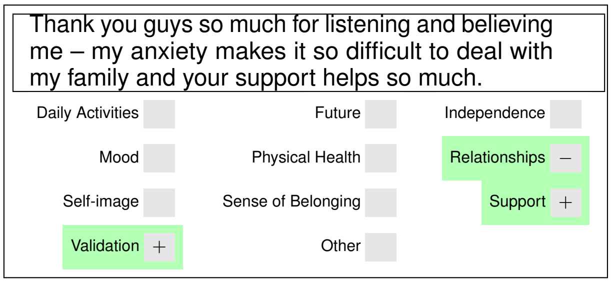

Figure 1: Annotated example of a Reddit post with annotated QoL aspects and sentiment labels in our corpus.

Automatically analyzing patient-centric reports of subjective QoL on social media potentially overcomes these issues. So far, however, they have been disregarded for QoL research, presumably because manually analyzing large quantities of social media posts is infeasible. To date, no resource exist that allows us to automatize the task. While prior work analyzes Twitter posts to gauge general QoL (Zivanovic et al., 2020) and to predict if a user's QoL is high or low (Sarma et al., 2019), these approaches do not capture the concrete aspects in which patients' QoL is affected.

Focusing on mental disorders, we are the first to leverage social media data for understanding health-related QoL and contribute QoLAS, a novel Reddit dataset annotated with 11 QoL aspects and their associated sentiment (POSITIVE, NEGATIVE, NEUTRAL, MIXED).[1] QoLAS covers eight mental disorders.

Figure 1 shows an example from the dataset. Here, the user describes how their anxiety negatively impacts family dynamics and expresses their happiness about the community (Reddit) support. Following our novel annotation, we label the QoL

[1]Find our annotations and the code to retrieve the Reddit posts here: https://www.uni-bamberg.de/en/nlproc/resources/qolas/.

aspects RELATIONSHIPS with a negative and SUPPORT and VALIDATION each with a POSITIVE sentiment label.

We answer the following research questions:

RQ1 (How) Do people discuss health-related QoL aspects on Reddit?

RQ2 Does aspect-based topic modeling reliably reflect QoL aspects?

RQ3 Does social media provide novel information on QoL aspects?

Our results show that Reddit users frequently discuss health-related QoL in their posts, the most frequent aspects being RELATIONSHIPS. Notably, the vast majority of QoL mentions on Reddit do not fit into any pre-existing aspect categories from health-related QoL research, highlighting how social media covers novel aspects relevant to patients' QoL. Our models reliably predict QoL mentions (.68 macro average F_1 score) and their sentiment (.55 macro average F_1 score). However, topic modeling does not appear to be the appropriate approach to identify individual, predefined aspects but provides an overview of topics under discussion. In a large-scale analysis of $\sim$56K sentences from Reddit posts discussing mental disorders, we show that social media is a valuable resource to broaden our understanding of subjective QoL aspects. We find that *studying* and *finances* are prevalent topics, showcasing new QoL dimensions which are not covered by existing QoL taxonomies.

2 Related Work

2.1 Social Media Health Mining

Social media health mining leverages computational methods to extract and analyze user-generated content related to medical conditions from social platforms (Klein et al., 2023). Under the objective of public health monitoring, Sarker et al. (2016) utilize social media for pharmacovigilance, i.e., analyzing mentions of adverse drug reactions. Further, social media has been employed to aid healthcare professionals in clinical decision-making and diagnosis (Sankhavara, 2018; Roller et al., 2022; Musen et al., 2021).

People frequently turn to social media as a safe space to discuss their medical journeys (Cohan et al., 2018). To this end, Murarka et al. (2021) develop a Reddit dataset to detect posts related to five mental disorders — depression, anxiety, bipolar disorder, ADHD (attention-deficit/hyperactivity disorder), and PTSD (post-traumatic stress disorder). Coppersmith et al. (2014) devise a dataset of diagnostic self-reports of mental disorders from social media. Similarly, Cohan et al. (2018) identify patterns of self-reported diagnoses and construct the self-reported mental health diagnoses data set. Jiang et al. (2020) explore linguistic markers to differentiate among mental disorders.

2.2 Quality of Life

Definition. The World Health Organization defines QoL as an "individuals' perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns" (World Health Organization, 2012, p.11).

It therefore constitutes a subjective, private state of a person (van Krugten et al., 2021; Eyl et al., 2018; Connell et al., 2014; Brazier et al., 2014; Karimi and Brazier, 2016).

Health-related QoL narrows QoL down to aspects relevant to a person's health (Yin et al., 2016). It is defined as referring to "how well a person functions in their life and his or her perceived well-being in physical, mental, and social domains of health" (Hays and Reeve, 2008). Health-related QoL, therefore, covers both aspects that can be observed from the outside, e.g., dressing oneself, employment status, walking/running, interactions with family/friends, and internal aspects such as one's subjective perception of pain, anxiety, depressive symptoms (Hays and Reeve, 2008).[2]

Measuring QoL. Generally, QoL measurements predominantly focus on physical aspects, for instance, walking, climbing stairs, bending, or kneeling (Ware and Sherbourne, 1992). To extend these variables to mental health related concepts, van Krugten et al. (2021) developed the Mental Health Quality of Life questionnaire (MHQoL), which covers self-image, independence, mood, relationships, daily activities, physical health, future.

We are not aware of any previous work in natural language processing (NLP) that focused on QoL in a medical sense on social media, however, there has been research on related concepts. Zivanovic et al. (2020) study people's overall QoL perception by assessing tweets focusing on topics such as transportation or parking. Sarma et al. (2019) assess the general level of health-related QoL on Twitter (QoL, i.e., high vs. low quality).

We extend this prior work with a dataset that

[2]We refer to health-related QoL using "QoL" in this paper.

enables us to detect health-related QoL aspects.

3 Corpus Creation and Analysis

We aim at understanding how Reddit users discuss the effects of mental health conditions on their quality of life. To this end, we create a corpus with annotations for QoL aspects and the associated sentiment polarity. Figure 1 shows an example.

3.1 Data Collection

Following Jiang et al. (2020), we focus on eight mental health conditions: anxiety, bipolar disorder (Bipolar), borderline personality disorder (BPD), depression, eating disorders, obsessive-compulsive disorder (OCD), post-traumatic stress disorder (PTSD), and schizophrenia. Based on a manual search and membership count, we select the following subreddits to collect texts from: r/Anxiety, r/bipolar, r/BPD, r/depression, r/EatingDisorders, r/OCD, r/ptsd, and r/schizophrenia.

We collect data from Reddit via the Python Reddit API Wrapper[3] and store the post text together with its title, author, post id, url, creation date, up/downvote score, number of comments, length of the post in token count, the name of the subreddit, and the information by the Reddit platform if a post is categorized as hot, new, or top. We exclude posts pinned by moderators (declaring rules and information for subreddits), posts that only contain images, and posts that do not have both a title and a main post text. For annotation, we sample the top 15 posts of the category hot and top, respectively for each mental health condition. This results in 240 posts for annotation.

3.2 Annotation

3.2.1 Annotation Task

The annotation consists of two subtasks: (1) Given a sentence from a Reddit post, annotators assign QoL aspect labels (see Table 1). (2) Annotators label the aspect-associated sentiment polarity. We now explain these steps in more detail.

QoL Aspects. Table 1 displays all possible QoL aspect labels. The first seven aspects (SELF-IMAGE, INDEPENDENCE, MOOD, RELATIONSHIPS, DAILY ACTIVITIES, PHYSICAL HEALTH, FUTURE) are taken from van Krugten et al. (2021). We further introduce the aspects of SUPPORT (the WHO declares the importance of support for people with mental health disorders[4]), SENSE OF BELONGING (report of people with mental health disorders perceiving sense of belonging as a self-reported QoL aspect (Connell et al., 2014)), and VALIDATION (Geller et al. (2021) find that validation is an important aspect for people with eating disorders). Finally, we introduce the aspect OTHER, allowing insight into QoL aspects that are not covered by our chosen label set.

In our study, annotators label each instance with one or more of the QoL aspects described in Table 1. For instance, in the sentence 'I am so disappointed that my family is not supporting me on my healing path', annotators would assign the QoL aspect SUPPORT. We provide the annotation guidelines in Appendix C.

Sentiment. Annotators label each aspect with its sentiment, i.e., one out of the labels POSITIVE, NEGATIVE, NEUTRAL or MIXED. In the example above, the SUPPORT is labeled as NEGATIVE.

3.2.2 Annotation Procedure

Setup. We split the Reddit posts into chunks to limit the cognitive complexity of parsing lengthy posts for the annotators. Each chunk consists of up to 7 sentences which is the median post length in the dataset. Annotators assign labels on the sentence level.

Environment. We create a custom annotation environment based on Google Sheets[5]. Each chunk is displayed individually. Annotators first decide if it contains a QoL assessment. If yes, the annotator extracts the relevant sentences and labels the aspects and their sentiment polarities. Annotators can reject chunks if they find them upsetting.

Annotators. We annotate 140 and 105 sentences in two iterations, employing three annotators. All annotators are aged 25–30 and have a background in computational linguistics and no medical training. Annotators A1 and A2 are female, A3 is male. A1 annotates the remaining 2,673 sentences.

3.2.3 Annotator Agreement

Evaluation metrics. We evaluate the annotations using two different metrics: average pairwise Cohen's κ and average pairwise inter-annotator F_1, where we regard one annotator's labels as the gold standard (Hripcsak and Rothschild, 2005).

[3]https://praw.readthedocs.io/en/stable/

[4]https://www.who.int/news-room/fact-sheets/detail/mental-disorders

[5]https://www.google.com/docs/about/

QoL Aspect	Abbr.	Explanation	Example
SELF-IMAGE.	S-I	How a person thinks about themselves (positively/negatively).	... I tried drawing, singing, playing instruments, **I'm just not good at anything**, and seeing people with actual talent just makes me mad...
INDEPENDENCE.	Ind.	How satisfied a person is with their level of INDEPENDENCE with respect to their freedom of choice, financial aspects, and (co-) decision-making.	... my mental illness has taken so much of me but the worst part is that **I feel so dependent on others** since on bad days I cannot even leave the house on my own ...
MOOD.	Mo.	Extend to which a person feels anxious, gloomy, or depressed.	... the heart pounding when you lay down, the twitching. . . **it just makes me so sad**...
RELATIONSHIPS.	Rel.	How satisfied a person is with their RELATIONSHIPS with their partners, children, family, or friends.	... Mid 20s, no job or any relevant experience at all. **I just leech off my parents while they pay for my meds and I couldn't do it without them**...
DAILY ACTIVITIES.	DA	How satisfied a person is with their daily activities with respect to work, study, household, or leisure activities.	... my anxiety makes it **hard for me to drive to work so sometimes I have to take a day off** just because I can't force myself to go there ...
PHYSICAL HEALTH.	PH	Any aspect that is related to a physical health problem of a person.	... then when I was 12 I developed bulimia. I'm 19 now and **my teeth are so fucked I cant chew anything at all**...
FUTURE.	Fut.	How optimistic/gloomy a person perceives their future.	... maybe it's because of the depression **I can't practice enough to be good at anything, and I know this will never change**...
SUPPORT.*	Sup.	Any aspect that does or does not make a person feel supported by friends, family, co-workers, people online, or people around them.	... EDIT: OMG you guys are the best. **the support I receive from you guys gives me the strength to keep pushing forward**...
SENSE OF BELONGING.*	SoB	Any aspect that makes the person feel like they belong to a specific group (friends, family, co-workers, people online, or people around them).	... since I joined this group I finally have the feeling that **I can share my experiences without being judged and honestly for the first time ever i don't feel like an outsider**...
VALIDATION.*	Val.	Any aspect that makes a person feel like their feelings or emotions are (in-) validated by friends, family, co-workers, people online, or people around them.	... the contamination OCD is strong today. People tell me to chill and stop disinfecting my hands every two seconds but **no one understands the pressure I have to do it and how I can't just 'chill' about it**...
OTHER.*	Oth.	Any aspect that is a QoL aspect but does not fit any of the other aspects.	... eat an orange in the shower. It helped me because **I have problems getting in, the temperature chance, the liquid, it's a lot of stimulus**. But the orange became my only concern...

Table 1: Taxonomy of QoL aspects annotated in QoLAS, following van Krugten et al. (2021). Aspects labeled with * are novel, meaning they have not been considered in the context of QoL for mental health so far.

	A1–A2		A1–A3		A2–A3		A1–A1	
	F_1	κ	F_1	κ	F_1	κ	F_1	κ
S-I	.60	.21	.71	.43	.63	.27	.80	.79
Ind.	.50	–.01	.49	–.01	.49	–.02	–	–
Mo.	.75	.50	.55	.15	.61	.25	.50	47
Rel.	.82	.65	.92	.84	.76	.53	1.00	1.00
DA	.60	.21	.55	.13	.75	.50	–	–
PH	.75	.50	.70	.41	.57	.16	1.00	1.00
Fut.	.49	–.03	.56	.13	.62	.24	–	–
Sup.	.78	.56	.69	.39	.77	.53	.67	.66
SoB	.49	–.01	.64	.27	.49	–.01	–	–
Val.	.49	–.02	.50	–.01	.64	.28	.67	.66
Oth.	.60	.21	.58	.20	.50	.09	.54	.85
¬QoL	.72	.43	.72	.45	.70	.40	.87	84

Table 2: F_1 and κ inter- and intra- agreement across labels. Hyphens indicate that no instances were annotated for a given label. Abbreviations are introduced in Table 1. ¬QoL indicates all sentences labeled with not containing a QoL mention.

	¬QoL		QoL	
	Sentences	Length	Sentences	Length
Anxiety	196	13.6	83	20.0
Bipolar	150	12.6	48	21.1
BPD	522	13.0	127	19.3
Depression	240	10.8	80	16.8
ED	243	13.0	70	19.6
OCD	252	15.2	32	21.0
PTSD	315	14.6	66	18.3
Schizophrenia	222	12.7	27	18.4
$\sum$	2140	13.23	533	19.18

Table 3: Numbers of labeled sentences across 8 mental disorders in QoLAS. The length is calculated as the mean average of tokens per sentence.

	Anxiety	Bipolar	BPD	Depression	ED	OCD	PTSD	Schizophrenia	Total
S-I	5	7	28	12	23	6	6	3	90
Ind.	2	0	1	2	0	0	0	1	6
Mo.	16	2	14	12	15	2	6	0	67
Rel.	12	16	36	7	12	2	12	0	97
DA	21	8	6	13	3	5	5	4	65
PH	4	1	3	10	4	2	1	1	26
Fut.	3	2	3	6	2	0	3	0	19
Sup.	13	9	9	3	7	1	2	2	46
SoB	4	7	14	5	2	1	1	1	35
Val.	4	2	10	2	5	5	9	3	40
Oth.	37	20	65	40	31	18	40	22	273
Total	117	74	189	111	104	42	45	17	748

Table 4: Counts of QoL aspect annotations across subreddits and QoL aspects.

	POSITIVE		NEGATIVE		NEUTRAL		MIXED	
	#	%	#	%	#	%	#	%
S-I	8	9	75	83	1	1	6	7
Ind.	3	50	3	50	0	0	0	0
Mo.	5	6	57	86	1	2	4	6
Rel.	2	2	71	73	1	1	23	24
DA	1	2	61	94	0	0	3	5
PH	2	7	25	89	0	0	1	4
Fut.	1	6	18	94	0	0	0	0
Sup.	17	38	16	35	1	2	12	25
SoB	7	20	19	54	2	6	7	20
Val.	3	8	31	78	1	2	5	12
Oth.	28	9	233	78	11	4	29	9
Total	72	10	575	77	18	2	83	11

Table 5: Distribution of QoL aspects across their sentiment labels.

Agreement. Table 2 shows the agreement scores for QoL aspects. The aspect RELATIONSHIPS shows the highest agreement, followed by SUPPORT. For SELF-IMAGE, MOOD, DAILY ACTIVITIES, and PHYSICAL HEALTH, the agreement is moderate. We find consistent results across annotators for 6 aspects (SELF-IMAGE, INDEPENDENCE, RELATIONSHIPS, FUTURE, SUPPORT, OTHER).

In contrast, the agreement varies between moderate for one annotator pair and slight for two pairs (MOOD, DAILY ACTIVITIES) and moderate for two pairs and slight for one (PHYSICAL HEALTH). Noteworthily, we find an agreement lower than chance for two annotator pairs for SENSE OF BELONGING and VALIDATION (the other being fair).

To ensure the consistency of the annotation approach used for the creation of our dataset, we further report the intra-annotator agreement of one annotator. A1 re-annotates 20 chunks 6 weeks after the first annotation. Table 2 shows the results. Overall, we observe a high agreement. For RELATIONSHIPS and PHYSICAL HEALTH, the agreement is perfect, while MOOD shows the lowest result.

3.3 (How) Do people discuss health-related QoL aspects on Reddit? (RQ1)

Corpus statistics. Table 3 presents the corpus statistics. The final corpus contains 2,140 sentences labeled with QoL aspect and sentiment. Out of those, 24.9% contain at least one QoL label. Sentences containing a QoL mention are consistently longer (mean average of tokens per sentence 19.18 for QoL, 13.23 for ¬QoL).

QoLAS contains 748 QoL labels in 533 sentences. From the selection of subreddits, those about BPD and eating disorders contain the most QoL aspects, schizophrenia the least. Texts from the subreddits corresponding to bipolar and OCD are the longest, and texts from depression are the shortest.

Which QoL aspects do users discuss on Reddit? We want to understand which aspects are relevant for people with mental disorders. Table 4 shows the distribution of QoL labels across conditions in our annotated corpus of Reddit sentences. The aspect OTHER is most frequent in QoLAS (273 mentions), followed by RELATIONSHIPS (97) and SELF-IMAGE (90). Importantly, the predominance of OTHER indicates that a substantial amount of aspects are not encapsulated within the confines of our pre-defined aspect labels. Generally, the distribution of aspects is imbalanced across the different mental health conditions and individual aspects. INDEPENDENCE (6) is a noticeable outlier.

How do users discuss QoL aspects on Reddit? Table 5 displays the distribution of QoL labels across all sentiment labels. The majority of aspects are labeled as NEGATIVE (77%), and the minority as NEUTRAL (2%). Notably, for INDEPENDENCE, we find an equal distribution of positive and negative labels and a mixed distribution of labels for SENSE OF BELONGING (20% positive, 54% negative, 6% neutral, 20% mixed). SUPPORT shows the least amount of negative sentiment labels (35%) and (except for the outlier INDEPENDENCE) the highest amount of positive sentiment labels (38%). This finding is in line with studies pointing out the positive influence of social support on mental health (Harandi et al., 2017; Turner and Brown, 2010).

Which QoL aspects co-occur? The majority

Aspect combinations	Frequency
RELATIONSHIPS, OTHER	33
MOOD, OTHER	27
DAILY ACTIVITIES, OTHER	25
SELF-IMAGE, OTHER	22
SELF-IMAGE, RELATIONSHIPS	20
VALIDATION, OTHER	14
PHYSICAL HEALTH, OTHER	12
FUTURE, OTHER	10
MOOD, RELATIONSHIPS	9
RELATIONSHIPS, SUPPORT	9

Table 6: Frequencies of top 10 QoL aspect combinations in QoLAS.

of QoL-related sentences contain exactly one aspect (61%). We find two aspects in 33% of QoL-related sentences and three or more aspects in 6%. To better understand which aspects co-occur, we analyze frequencies of pairs in sentences in Table 6. The most frequent combination is the tuple (RELATIONSHIPS, OTHER) (33) followed by (MOOD, OTHER) (27). Interestingly, the frequency of the tuple (SELF-IMAGE, RELATIONSHIPS) (20) suggests a relation between one's own or perceived self-image within a relationship. The aspect OTHER appears in all four most frequent tuples. Overall, within the top 10 QoL aspect combinations, we find the aspect of OTHER in combination with seven aspects (RELATIONSHIPS, MOOD, DAILY ACTIVITIES, SELF-IMAGE, VALIDATION, PHYSICAL HEALTH, FUTURE).

Our analysis of the QoL label distribution in QoLAS highlights the importance of the aspect OTHER. This indicates that our QoL label set, based on the MHQoL, only partially covers the variety of QoL aspects. This assumption is further supported by the QoL aspect combinations, where we find that most QoL aspects appear in a sentence where in addition to a QoL aspect from the MHQoL the aspect OTHER is also represented.

4 Experiments

We investigate *how reliably we can detect QoL aspects and their sentiment polarity automatically* (RQ2). To this end, we build a pipeline that consists of three modules: First, we train a classifier that detects if a sentence contains a QoL mention. Second, we leverage a topic model to detect the QoL aspect. Finally, we employ a sentiment classifier to predict the sentiment polarity. Figure 2 shows a depiction of the pipeline. We describe each module in the following.

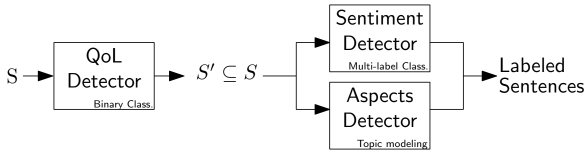

Figure 2: Automatic QoL aspect and sentiment detection. We simplify the aspect-based sentiment analysis to an independent sentence classification method.

4.1 Methods

We train and evaluate the models of the QoL detection and sentiment classification modules using an 80/20 train/test split of QoLAS. We provide implementation details for all modules in Appendix B.

QoL detection. In the first module, we determine if a text sequence contains a QoL aspect, based on two random forest and four pre-trained language models.

Specifically, we utilize two Random Forests (Breiman, 2001) as baselines, namely RF_{base} and $\text{RF}_{\text{balanced}}$ and further fine-tune four transformer models, specifically BERT (Devlin et al., 2019), RoBERTa (Liu et al., 2019), DistilBERT (Sanh et al., 2019) and SocBERT (Guo and Sarker, 2023). We provide the implementation details for all models and individual modules of the pipeline in Appendix B.

Aspect detection. To predict specific QoL aspects, we employ a topic model, namely BERTopic (Grootendorst, 2022), a transformer-based clustering approach. BERTopic is a method that uses transformer embeddings, capturing the contextual meaning of words, which is advantageous over other topic modeling options such as LDA. We tune the underlying model to find topics semantically similar to the QoL aspect labels.

Sentiment classification. We fine-tune BERT on QoLAS for sentiment polarity prediction with the target classes POSITIVE, NEGATIVE, NEUTRAL, and MIXED.

4.2 Evaluation

We evaluate the performance of all models on the held-out test set of QoLAS and report precision, recall, and F_1.

We frame the QoL-related aspect detection as unsupervised modeling to enable our system to uncover yet unknown aspects. With this perspective in mind, we evaluate the topic modeling as an open-domain aspect recognition process. As such, we annotate 100 sentences with 18 topics identified ac-

	¬QoL			QoL			Macro Avg.		
	P	R	F_1	P	R	F_1	P	R	F_1
RF_{base}	.80	.98	.89	.37	.04	.07	.59	.51	.48
$RF_{balanced}$	.91	.70	.79	.37	.71	.49	.64	.71	.64
BERT	.90	.88	.89	.48	.53	.50	.67	.73	.68
RoBERTa	.91	.85	.88	.46	.58	.50	.66	.73	.65
DistilBERT	.89	.88	.88	.47	.50	.47	.68	.69	.68
SocBERT	.91	.86	.87	.46	.59	.51	.66	.66	.65

Table 7: Precision, recall, and F_1 results of the baseline (RF) and different transformer models on the sentence-level QoL classification.

cording to a manual inspection of an independent development corpus of QoL-related sentences. We provide the topics in the Appendix in Section A. The evaluation therefore constitutes a clustering-based evaluation, for which we employ the adjusted rand index (ARI) (ARI, Hubert and Arabie, 1985) which calculates the similarity between clustering while correcting for cluster similarity by chance. ARI scores range from -1 to 1, where 1 indicates perfect agreement and negative values indicate an agreement worse than chance.

4.3 Results

4.3.1 QoL Detection

Table 7 shows the model performance in the QoL detection module. Overall, transformer models outperform the baseline models RF_{base} and $RF_{balanced}$. On average, BERT and DistilBERT are the most robust models (.68 macro avg. F_1, respectively). Notably, for the target class (QoL), SocBERT performs best across all models.

4.3.2 Aspect Detection

The best BERTopic model achieves an ARI score of 0.16 on the gold standard.

4.3.3 Sentiment Classification

Table 8 shows the performance of the sentiment classification module. Considering the macro average performance across all classes, the BERT model fails to beat the majority baseline (.55 F_1 vs. .73 F_1).

The model performs most reliably for the POSITIVE and NEGATIVE class (.70 F_1 and .92 F_1, respectively). For the MIXED class the model achieves an F_1-score of .56. For the NEUTRAL instances, the classifier cannot predict a single correct instance, due to the class imbalance in the data set.

	Precision	Recall	F_1
Positive	.78	.64	.70
Negative	.88	.95	.92
Neutral	.00	.00	.00
Mixed	.58	.54	.56
Macro Avg.	.56	.53	.55
Majority Baseline			.73

Table 8: Results of the BERT model on the sentiment classification task on QoLAS sentences.

Overall, our results show a moderate performance of transformer models on the QoL detection task (.68 macro avg. F_1) and sentiment polarity prediction (.55 macro avg. F_1). Correctly predicting the specific QoL aspects using a topic modeling approach, specifically BERTopic, remains challenging, indicating that topic modeling is not the most fitting approach for our task.

5 Does Social Media Provide Novel QoL Aspects?

We hypothesize that quality of life aspects that people discuss on social media may be even more fine-grained than the aspect labels in the QoLAS dataset and hold entirely novel QoL aspects (RQ3). To investigate this, we conduct two qualitative analyses on a large set of mental health-related Reddit posts.

5.1 Data

We collect 125,994 posts from Reddit. Analogous to the data crawling and filtering in Section 3.1, we collect posts associated with the same 8 mental disorders as in QoLAS. We split them into sentences and employ the best performing QoL detection model, namely BERT[6], to identify instances discussing QoL mentions. This provides us with 55,920 sentences. Subsequently, we employ the topic model to obtain prevalent topic clusters within this data.[7]

5.2 Analysis

We aim at understanding how the automatically identified topic clusters relate to the established set of QoL aspects we use in QoLAS. Therefore,

[6] BERT shows superior performance to DistilBERT on the target class QoL (.50 F_1 and .47 F_1).

[7] We acknowledge that the topic model did not show a robust performance to detect manual annotations of aspects. Nevertheless, it provides a meaningful way to aggregate information in the corpus we study here, as we are not aiming to detect this pre-defined set of aspects.

QoL Aspect	Topic
SELF-IMAGE	–
INDEPENDENCE	–
MOOD	emotional_anger_rage, life_depression_hate, exhausted_tired_unmotivated, emotional_anger_rage, depressive_depressants_depresh, apprehensions_fears_paranioa, cried_sobbed_sobbing
RELATIONSHIPS	friendless_friendships_friends, dating_date_relationships, lonely_loneliness_horny
DAILY ACTIVITIES	sleep_asleep_insomnia
PHYSICAL HEALTH	diet_underweight_overweight
FUTURE	–
SUPPORT	supportive_insecurely_scold
SENSE OF BELONGING	–
VALIDATION	–
OTHER	hunger_hungry_appetite, contamination_compulsions-_contaminating, panicky_hyperventilating_panic, neurosis_bpad_bdp, grades_studying_study, suicidal_suicide_kill, therapy_therapist_counseling, savings_finances_paycheck

Table 9: Top 20 topics (of 55,920 sentences), generated by BERTopic, in comparison to 11 QoL aspect labels.

we manually map the top 20 topic clusters generated by the topic model to the QoLAS label set. Any topic that is not connected to the labels, we attribute to the aspect OTHER. Based on this, we analyze (a) if the topics may be more fine-grained compared to established QoL labels and (b) which novel topics are relevant in mental health-related online discussions that are not covered by current QoL research.

Table 9 shows the result of the mapping. Overall, we are able to map half of the automatically generated clusters to the QoLAS label set. For the QoL aspects of DAILY ACTIVITIES, PHYSICAL HEALTH, and SUPPORT we find one corresponding topic each. For the QoL aspect RELATIONSHIPS and MOOD, we find 3 and 7 corresponding topics, respectively. For SELF-IMAGE, INDEPENDENCE, FUTURE, SENSE OF BELONGING, and VALIDATION we do not find a corresponding topic. Notably, all aspects without corresponding topics are relatively abstract. We hypothesize that these concepts tend to be non-propositional and expressed differently by individual users. The topic model therefore may not be able to capture them, while aspects such as MOOD or RELATIONSHIPS are more concrete.

We map 7 topics to the aspect OTHER. This indicates that they are not related to any of the established QoL aspects. The topics in this category are diverse, with 4 out of 8 topics appear to be associated with specific medical conditions or symptoms thereof (contamination, compulsions, contaminating; panicky, hyperventilating, panic; neurosis, bpad, bdp; suicidal, suicide, kill). The other topics are more general, e.g., centered around studying and finances, indicating that mental disorders affect core aspects of people's lives.

Our analysis indicates that QoL dimensions go beyond the aspects that are covered by current QoL taxonomies, emphasizing the potential of accessing user-generated data for mental health research and knowledge discovery.

6 Conclusion

To address the limitations of small-scale QoL assessments, we leverage Reddit posts to extract health-related QoL aspects and extend our understanding to novel relevant aspects. We contribute QoLAS, the first dataset to model health-related QoL aspects for mental disorders automatically. We show that Reddit posts provide detailed medical accounts in which users discuss a multitude of health-related QoL aspects. Using open-domain aspect-based sentiment analysis, we are able to reliably detect QoL discussions and their sentiment. However, topic modeling struggles to identify individual QoL aspects within our dataset. This leads to important future work, namely to explore other methods to detect such detailed properties of QoL. Importantly, we find that a substantial number of QoL mentions go beyond the established taxonomy we use to label social media posts. This is true for our gold-labeled dataset, and holds for a large-scale analysis of Reddit posts. For those instances, it is crucial for future work to obtain a detailed understanding of the themes and topics that emerge. They have the potential to inform medical practitioners, particularly for underrepresented demographics, or the effects of rare symptoms that are out of the scope of existing resources.

Limitations

With respect to the annotation of QoL aspects in posts, we obtain a robust agreement for the target class of QoL documents, however, we acknowledge that the agreement scores for the aspects vary. The annotation is a challenging task due to the subjective nature of QoL aspects. However, examining the robust F_1-scores we consider the agreement acceptable. Further, we maintain confidence in the quality of the annotations despite the corpus being labeled by a single annotator. While more annotators might account for individual biases and errors, one well-trained annotator strictly following carefully constructed guidelines can produce high-quality annotations. This can be seen in our high intra-annotator agreement in Table 2.

While our annotators are non-medical experts, we consulted medical experts, who specialized in quality-of-life research, during the conceptualization of the QoL aspect annotation, ensuring the correctness of our annotation approach from a medical perspective.

Ethical Considerations

Studying QoL aspects of people with mental disorders has to be done carefully to prevent potential harm. We make use of posts from subreddits that are created for specific mental disorders. We are assuming that people writing posts there are (self-)diagnosed with a given mental disorder. With our QoL aspect annotation, we, to a certain extend, infer the well-being of specific users, which can be perceived as upsetting.

Therefore, it is crucial to use respectful language when describing these posts or displaying results to avoid perpetuating stigma around mental disorders.

We ensured the annotator's safety and mental well-being by warning them about possibly disturbing content. Annotators were instructed to only annotate posts that do not make them feel uncomfortable in any way and to take breaks if needed. In addition, they had the option to reject each chunk of text individually. We note that no annotator made use of this option.

The data we collected in our study is solely used for academic purposes. We strictly follow Reddit's guidelines on data distribution and do not publish the data itself. Instead, to still enable follow-up research, we provide a script[8] that allows researchers to collect the data from Reddit and match it to our annotations. We are aware that our choice of using Reddit posts complicates access to our created corpus, however, this type of data is relevant for our research purposes.

[8] https://www.uni-bamberg.de/en/nlproc/resources/qolas/

Acknowledgments

This work has partially been supported by the EMCONA project (The Interplay of Emotions and Convincingness in Argument Mining for NLP) and the FIBISS project (Automatic Fact Checking for Biomedical Information in Social Media and Scientific Literature), both funded by the German Research Council (DFG, project numbers: KL 2869/5-1 and KL 2869/12-1).

References

John Brazier, Janice Connell, Diana Papaioannou, Clara Mukuria, Brendan Mulhern, Tessa Peasgood, Myfawnwy Lloyd Jones, Suzy Paisley, Alicia O'Cathain, Michael Barkham, Martin Knapp, Sarah Byford, Simon Gilbody, and Glenys Parry. 2014. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. *Health Technol Assess*, 18(34):1–188.

Leo Breiman. 2001. Random forests. *Machine Learning 45*, page 5–32.

Arman Cohan, Bart Desmet, Andrew Yates, Luca Soldaini, Sean MacAvaney, and Nazli Goharian. 2018. SMHD: a large-scale resource for exploring online language usage for multiple mental health conditions. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1485–1497, Santa Fe, New Mexico, USA. Association for Computational Linguistics.

Janice Connell, Alicia O'Cathain, and John Brazier. 2014. Measuring quality of life in mental health: are we asking the right questions? *Soc Sci Med*, 120:12–20.

Glen Coppersmith, Mark Dredze, and Craig Harman. 2014. Quantifying mental health signals in Twitter. In *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*, pages 51–60, Baltimore, Maryland, USA. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Ruth Elisa Eyl, Kun Xie, Lena Koch-Gallenkamp, Hermann Brenner, and Volker Arndt. 2018. Quality of life and physical activity in long-term (≥5 years post-diagnosis) colorectal cancer survivors - systematic review. *Health Qual Life Outcomes*, 16(1):112.

Josie Geller, A. Fernandes, S. Srikameswaran, R. Pullmer, and S. Marshall. 2021. The power of feeling seen: perspectives of individuals with eating disorders on receiving validation. *Journal of Eating Disorders*, 9(1):149.

Maarten R. Grootendorst. 2022. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *ArXiv*, abs/2203.05794.

Yuting Guo and Abeed Sarker. 2023. SocBERT: A pretrained model for social media text. In *The Fourth Workshop on Insights from Negative Results in NLP*, pages 45–52, Dubrovnik, Croatia. Association for Computational Linguistics.

Tayebeh Fasihi Harandi, Maryam Mohammad Taghinasab, and Tayebeh Dehghan Nayeri. 2017. The correlation of social support with mental health: A meta-analysis. *Electron Physician*, 9(9):5212–5222.

Ron D. Hays and Bryce B. Reeve. 2008. Measurement and modeling of health-related quality of life. *International Encyclopedia of Public Health*, pages 241–252.

George Hripcsak and Adam S. Rothschild. 2005. Agreement, the f-measure, and reliability in information retrieval. *Journal of the American Medical Informatics Association: JAMIA*, 12(3):296–298.

Lawrence Hubert and Phipps Arabie. 1985. Comparing partitions. *Journal of Classification*, 2(1):193–218.

Zhengping Jiang, Sarah Ita Levitan, Jonathan Zomick, and Julia Hirschberg. 2020. Detection of mental health from Reddit via deep contextualized representations. In *Proceedings of the 11th International Workshop on Health Text Mining and Information Analysis*, pages 147–156, Online. Association for Computational Linguistics.

Milad Karimi and John Brazier. 2016. Health, health-related quality of life, and quality of life: What is the difference? *PharmacoEconomics*, 34(7):645–649.

Ari Z. Klein, Juan M. Banda, Yuting Guo, Ana Lucia Schmidt, Dongfang Xu, Jesus Ivan Flores Amaro, Raul Rodriguez-Esteban, A. Sarker, and Graciela Gonzalez-Hernandez. 2023. Overview of the 8th social media mining for health applications (#smm4h) shared tasks at the amia 2023 annual symposium. *medRxiv*.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *ArXiv*, abs/1907.11692.

Ankit Murarka, Balaji Radhakrishnan, and Sushma Ravichandran. 2021. Classification of mental illnesses on social media using RoBERTa. In *Proceedings of the 12th International Workshop on Health Text Mining and Information Analysis*, pages 59–68, online. Association for Computational Linguistics.

Mark A. Musen, Blackford Middleton, and Robert A. Greenes. 2021. Clinical decision-support systems. In Edward H. Shortliffe and James J. Cimino, editors, *Biomedical Informatics*. Springer, Cham.

Roland Roller, Aljoscha Burchardt, Nils Feldhus, Laura Seiffe, Klemens Budde, Simon Ronicke, and Bilgin Osmanodja. 2022. An annotated corpus of textual explanations for clinical decision support. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 2317–2326, Marseille, France. European Language Resources Association.

Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. 2019. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *ArXiv*, abs/1910.01108.

Jainisha Sankhavara. 2018. Biomedical document retrieval for clinical decision support system. In *Proceedings of ACL 2018, Student Research Workshop*, pages 84–90, Melbourne, Australia. Association for Computational Linguistics.

Abeed Sarker, Karen O'Connor, Rachel Ginn, Matthew Scotch, Karen Smith, Dan Malone, and Graciela Gonzalez. 2016. Social media mining for toxicovigilance: Automatic monitoring of prescription medication abuse from twitter. *Drug Safety*, 39(3):231–240.

Karthik V. Sarma, Brennan M. R. Spiegel, Mark W. Reid, Shawn Chen, Raina M. Merchant, Emily Seltzer, and Corey W. Arnold. 2019. Estimating the Health-Related Quality of Life of Twitter Users Using Semantic Processing. *Stud Health Technol Inform*, 264:1065–1069.

R. Jay Turner and Robyn Brown. 2010. Social support and mental health. *A Handbook for the Study of Mental Health: Social Contexts, Theories, and Systems*, pages 200–212.

Frédérique C. W. van Krugten, Jan J. V. Busschbach, Matthijs M. Versteegh, Leona Hakkaart van Roijen, and Werner B. F. Brouwer. 2021. The mental health quality of life questionnaire (mhqol): development and first psychometric evaluation of a new measure to assess quality of life in people with mental health problems. *Quality of Life Research*, 31:633 – 643.

John E. Ware and Cathy Donald Sherbourne. 1992. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*, 30(6):473–483.

World Health Organization. 2012. Programme on mental health: WHOQOL user manual, 2012 revision.

Shaoman Yin, Rashid Njai, Lawrence Barker, Paul Z. Siegel, and Youlian Liao. 2016. Summarizing health-related quality of life (hrqol): development and testing of a one-factor model. *Population Health Metrics*, 14:22.

Slavica Zivanovic, Javier Martinez, and Jeroen Verplanke. 2020. Capturing and mapping quality of life using twitter data. *GeoJournal*, 85(1):237–255.

A Appendix: Topic Model Evaluation

We report details on the gold-standard (100 manually labeled topics, see Section 4.2 for the evaluation of BERTopic in Table 10. Table 11 provides the output of topics and numbers of sentences of our tuned BERTopic model on the same 100 sentences.

Topic	# Sent.	Topic	# Sent.
Eating	7	Hope	11
Sleeping	6	Everyday	8
Relationships	8	Worry	4
Emotions	5	Self-image	3
Support	5	Lost motivation	10
Alone	2	Finances	2
Existence	3	Exhaustion	5
Sadness	9	Physical health	3
Understanding	5	Symptom	4

Table 10: Distribition of labels of 18 topics (identified according to a manual inspection of an independent development corpus of QoL-related sentences) for 100 sentences.

Topic	# Sent.
stress_worry_worried	19
relationships_abandonment_love	6
bulimia_eating_appetite	11
exhausted_exhausting_relax	6
life_hope_happ	10
happy_grateful_treatment	6
sleep_bed_wake	9
trust_stresses_gang	5
support_die_suffering	8
feel_feels_exist	5
overwhelmed_crying_emotional	6

Table 11: Output of BERTopic on the gold standard.

B Appendix: Implementation Details

In the following, we provide details on the implementation of all modules of our pipeline.

Aspect Detection. We implement the Random Forest Classifier using the scikit-learn implementation[9] and the Balanced Random Forest Classifier, using the imbalanced-learn implementation [10]. We use the default settings for both classifiers. We set the random state to 42.

For all transformer models, we use the respective model's PyTorch implementation from HuggingFace. For BERT, we use the bert-base-uncased model[11], for RoBERTa we use the xlm-roberta-base model[12], for distilBERT we use the distilbert-base-uncased model[13], and lastly, for SocBERT, we use the SocBERT-base model[14]. We fine-tune the models using the AdamW optimizer. We use a learning rate of $3{\cdot}10^{-5}$, batch size of 16, and 20 epochs.

BERTopic. For fine-tuning BERTopic, we follow `https://maartengr.github.io/BERTopic/index.html`(Grootendorst, 2022). We experiment with various hyperparameter settings to identify the most robust setup. The best setting (achieving .16 ARI) is: the all-mpnet-base-v2 as the embedding model, the CountVectorizer with English stop words, UMAP with a local neighborhood of 3 and dimension of space to embed into of 3, the HDBSCAN model with a cluster size of 3, the KeyBertInspired for the topic representation, and the number of topics are set to be adjusted automatically.

Sentiment Classification. We implement the bert-base-uncased model using its PyTorch implementation from HuggingFace [15]. We fine-tune the model for sentiment classification using the AdamW optimizer with a learning rate of $2{\cdot}10^{-5}$. We use a batch size of 8 and 6 epochs.

[9] `https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html`

[10] `https://imbalanced-learn.org/stable/references/generated/imblearn.ensemble.BalancedRandomForestClassifier.html`

[11] `https://huggingface.co/bert-base-uncased`

[12] `https://huggingface.co/xlm-roberta-base`

[13] `https://huggingface.co/distilbert-base-uncased`

[14] `https://huggingface.co/sarkerlab/SocBERT-base`

[15] `https://huggingface.co/bert-base-uncased`

C Appendix: Annotation Guidelines

Note that from here on, we display the annotation guidelines exactly as they have been shown to the annotators.

1. Instructions

1. Read over this document.
2. Open the provided link that will open the Google Spreadsheet.
3. In the Google Sheets file, go to Extensions (Erweiterungen) and click on "Apps Script". This opens a new window with the script for the annotation. Click on Run (Ausführen). Do not close this window, but go back to the Google Sheets and follow the instructions that pop up there.

2. The Task

We annotate Quality of Life (QOL) in mental health Subreddits (Anxiety, Depression, Schizophrenia, PTSD, OCD, Eating Disorders, Bipolar Disorder, and BPD). We are interested in the aspect and the sentiment of QOL aspects. In addition, we want to know if the author of a post is diagnosed with a condition or not.
In the following, you get detailed information about and instructions for the annotation.

3. QOL Definition

The WHO defines QOL to be the **"individuals' perceptions of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns"**
In addition:

- QOL aspects are subjective
- QOL aspects vary for different medical conditions and individuals

4. QOL Annotation: Aspects and Sentiment

4.1 Aspects

We are interested in the specific aspect of a given QOL aspect. The options are: selfimage, independence, mood, relationships, daily activities, physical health, future, support, sense of belonging, validation, and other. Note that some of the aspects can be overlapping (e.g. in the table: the example QOL aspect for independence could be both, independence and daily activities). Since this is often times the case, you are encouraged to annotate multiple aspects for the same QOL aspect, if it fulfills multiple aspects. The following table displays short descriptions of the aspects as well as an example for each aspect. You are encouraged to assign **'Other'**, if one or more of the following criteria holds:

- The QOL does not fit any of the other aspects.
- You can come up with an aspect that fits the QOL better than the given aspects.
- The QOL does only partly fit one of the aspects.

If in doubt, rather assign one of the descriptive aspects and **'Other'** instead of only the descriptive aspect(s).

Text	Aspect
... I tried drawing, singing, playing instruments, **I'm just not good at anything**, and seeing people with actual talent just makes me mad...	**Self-image.** I think positively/negatively about myself.
... my mental illness has taken so much of me but the worst part is that **I feel so dependent on others** since on bad days I cannot even leave the house on my own ...	**Independence.** I am very (dis-)satisfied with my level of INDEPENDENCE. (Freedom of choice, financial, co-decision making)
... the heart pounding when you lay down, the twitching... **it just makes me so sad**...	**Mood.** I (do not) feel anxious, gloomy, or depressed.
... Mid 20s, no job or any relevant experience at all. **I just leech off my parents while they pay for my meds and I couldn't do it without them**...	**Relationships.** I am very (dis-)satisfied with my relationships (Partner, children, family, friends)
... my anxiety makes it **hard for me to drive to work so sometimes I have to take a day off** just because I can't force myself to go there ...	**Daily Activities.** I am very (dis-)satisfied with my daily activities. (Work, study, household, leisure activities)
... then when I was 12 I developed bulimia. I'm 19 now and **my teeth are so fucked I cant chew anything at all**...	**Physical Health** I have no/many physical health problems.
... maybe it's because of the depression **I can't practice enough to be good at anything, and I know this will never change**...	**Future.** I am very optimistic/gloomy about my future.
... EDIT: OMG you guys are the best. **the support I receive from you guys gives me the strength to keep pushing forward**...	**Support.** I (do not) feel supported by my friends, family, co-workers, people online, people around me.
... since I joined this group I finally have the feeling that **I can share my experiences without being judged and honestly for the first time ever i don't feel like an outsider**...	**Sense of Belonging.** I feel like I belong to a specific group (friends, family, co-workers, people online, people around me).
... the contamination OCD is strong today. People tell me to chill and stop disinfecting my hands every two seconds but **no one understands the pressure I have to do it and how I can't just 'chill' about it**...	**Validation.** I (do not) feel validated by my friends, family, co-workers, people online, people around me.
... eat an orange in the shower. It helped me because **I have problems getting in, the temperature chance, the liquid, it's a lot of stimulus**. But the orange became my only concern... ... i wish i would have been better at visiting my grandparents who raised me, i wish i would've/could've answered my grandmas calls before she was gone. **i wish so bad this disorder didn't do the things it does to me because i live everyday of my life full of guilt**...	**Other.** Anything that is a QOL but does not fit any of the given aspects. Here, for instance, hygiene, regret, guilt.

4.2 Sentiment Label

In addition, we are interested in the sentiment of the specific QOL aspect. The options are: positive, negative, neutral, mixed. Find examples in the table below.

Text	Sentiment
... EDIT: OMG you guys are the best. **the support I receive from you guys gives me the strength to keep pushing forward**...	Positive (Aspect: Support)
... the heart pounding when you lay down, the twitching... **it just makes me so sad**...	Negative (Aspect: Mood)
... this group gives me the opportunity to **exchange experiences and coping strategies without necessarily impacting my overall emotional state. It's helpful for me to hear different perspectives**...	Neutral (Aspect: Sense of Belonging, Other)

... thank you for all the comments **they make me sad because we are all suffering but we will heal together I promise** ...	Mixed (Aspect: Mood, Support)

5. QOL Annotation: More Examples

In the following, you can find more examples of gold QOL annotations. Since we are annotating on a sentence level, there can be multiple sentences that contain QOL aspects (see example 2). In that case, you first annotate the QOL aspect(s) in the first sentence, then the aspect(s) in the second sentence, and so on. It might also happen that there is more than one QOL aspect in one sentence. In that case, you should just go over all of the aspects and annotate all of them with their respective sentiment. In the last table, you find some examples of text that do not contain any QOL aspects.

5.1 More QoL examples

Text	Aspect and Sentiment
... I question everything now about who I was and the people I trusted. **I feel so lost and alone and confused.** It's not common at my age...	Mood, Negative
... How did affect the other person? How did your response affect you? How can you respond better next time? I noticed I had 3 major symptoms that occurred more frequently than others. **Fear of abandonment.** **Unstable relationships.** **Unclear sense of self.**...	Future, Negative; Independence, Negative Relationships, Negative Self-image, Negative
... Yesterday I wrote about my experiences. I love you all. Edit: **Thank you so much for all of your comments they make me feel understood** and **help me to manage the days where I cannot get out of bed and make myself food.**...	Validation, Positive; Sense of Belonging, Positive Mood, Negative; Daily Activities, Negative

5.2 No QOL aspects

Text	Explanation
I finally booked an appointment with my GP (doctor). I'm trying to make an effort to get better. I was wondering what I should tell, how much should I open up with him, and if I should mention my suicidal thoughts. How does the process work?	Suicidal thoughts are a symptom of depression.

I'm back to normal, just like that. I'm queing for another match of my videogame. BPD is fucking real if any of you any had any doubt about it. From one second to another I became suicidal, guilty, angry, paranoid and delussional.	The person is describing symptoms of their illness, they are not talking about how it affects their QOL.
Just got a job at McDonald's. 40 + hours a week. Wish me luck!	No connection to how this affects the person's QOL.

5.3 QOL vs. Symptoms

There are cases in which it is difficult to decide whether a description you read is a QOL aspect or a symptom of the medical condition. In some cases, only a medical expert would be able to make that decision. Therefore, if you are unsure, rather assign a QOL aspect. Here are some considerations:

- Some symptoms and QOL aspects can impact an individual's QOL: assign QOL
- A description of feeling disconnected from reality can be a symptom of certain mental disorders but if it is not explicitly stated how that affects the QOL of an individual, treat it as a symptom (i.e. do not assign a QOL aspect)
- Consider the context: if feeling disconnected from reality is causing an individual to having trouble working out, and, therefore, having back pain, it is a QOL aspect

6. Annotation: Workflow

1. You will see a post or an excerpt from a post from Reddit and the corresponding subreddit (all subreddits are medical conditions, such as anxiety, ptsd, depression, etc.). Read over it. If the content makes you uncomfortable, you can directly discard the post and go to the next one. Please consider your own well-being and only annotate a post, if you are sure that it does not trigger you in any way.
2. Depending on your previous answer, you will either get a new post, so you repeat step 1, or you decided to annotate the post, which leads to the following: You will see a post, a title, and the condition (the subreddit) and you have to decide whether the document does or does not contain a mention of a QOL aspect. If you decide that there is no QOL aspect in the document, you will get a new post and start with step 1 again. If there is a QOL aspect, after clicking yes, you will proceed with step 3.
3. You will see the post again and have to decide if you think that the person who wrote the post is officially (medically) diagnosed with the condition that is displayed. Give a rating from 1-5 (1 means the person is most likely not diagnosed, and 5 means the person is most likely diagnosed). If the post explicitly states that the person is or is not diagnosed, assign a 5 or 1 respectively. If you have a strong intuition that the person is or is not diagnosed, without an explicit statement about it, assign a 4 or 1 respectively. If you have no justified assumption that the person is or is not diagnosed, assign a 3.
4. Now you are asked to select all the sentences of the post that do contain a QOL aspect. Type their number in and separate it with commas. You will go over all of the sentences you selected individually in the next step.
5. Now you are asked to provide the aspect(s) of the QOL aspect(s) in the text and the respective sentiment (e.g. Mood, Negative). The options for the sentiment of the aspect are positive, negative, neutral, or mixed (+, -, n, m) Be aware you might find multiple aspects in one sentence. Annotate all of them (e.g. Mood, Negative; Relationships, Mixed). If none of the aspects fit, but you are sure that there is a QOL aspect in the sentence, assign the aspect 'other'. Now you have successfully annotated all of the QOL aspects in one sentence. You will proceed to the next question.

6. Unfortunately, the Google Apps Script in which you are annotating is limited to a runtime of 6 minutes. Therefore, you will see a reminder to restart the script before the next post gets displayed. If you see the reminder, switch to the Apps Script window and click on “Stop” (Beenden) and then on “Run” (Ausführen) again. Then, return to the sheets window, click on ok on the reminder, and continue with the next post.

7. Additional Notes

If you want to take a break or are done with your annotation, ideally you would click on the x in the displayed box of the first question, if the post is too triggering. When you come back to the annotation, you will then be able to start right where you left off. You just need to click “Run” in the Script again.

Please be aware that it is not possible to change one of your answers or go back to one question with the provided script. If you accidentally clicked something wrong, you can look into the sheets file and see if you can easily change the value manually. If not, please take note of what happened, write down the ID of the post and send it to me.